A DATA SONIFICATION APPROACH TO COGNITIVE STATE IDENTIFICATION

Imanol Gomez

Univeristat Pompeu Fabra
Music Tecnology Group
Tanger 122-138
08018 Barcelona, Spain
imanol.gomez01@estudiant.upf.edu

Rafael Ramirez

Univeristat Pompeu Fabra
Music Tecnology Group
Tanger 122-138
08018 Barcelona, Spain
rafael.ramirez@upf.edu

ABSTRACT

The study of human brain functions has dramatically increased greatly due to the advent of Functional Magnetic Resonance Imaging (fMRI), arguably the best technique for observing human brain activity that is currently available. However, fMRI techniques produce extremely high dimensional, sparse and noisy data which is difficult to visualize, monitor and analyze. In this paper, we propose two different sonification approaches to monitor fMRI data. The goal of the resulting fMRI data sonification system is to allow the auditory identification of cognitive states produced by different stimuli. The system consists of a feature selection component and a sonification engine. We explore different feature selection methods and sonification strategies. As a case study, we apply our system to the identification of cognitive states produced by volume accented and duration accented rhythmic stimuli.

1. INTRODUCTION

The human brain is an extremely complex information processing system and the understanding of most of its functions is still a major challenge. Many techniques have been developed to detect and measure neural activity in humans (e.g. EEG, fMRI, CAT) and various methods have been proposed for analyzing the resulting data. In particular, Functional Magnetic Resonance Imaging (fMRI) has been used extensively to test hypothesis regarding the location of activation for different brain functions. However, fMRI provides extremely high dimensional, sparse and noisy data which is difficult to visualize, monitor and analyze.

The goal of exploratory data analysis is to render high dimensional data in such a way that we can use our natural pattern recognition capabilities in order to search for regularities and structures. This approach has mainly focused on human visual capabilities. Many visualization techniques have been developed such as Self-Organizing Maps [1, 2], Multidimensional Scaling [3] and Projection Pursuit [4] which creates low-dimensional imaging of the original data.

Motivated by the acknowledged human capacity to make accurate and rapid processing and discrimination of sounds, in this paper we investigate human auditory perception for exploring and analyzing fMRI data. In particular, we propose a sonification approach to monitor and exploring fMRI data. Our goal is to allow the auditory identification of cognitive states produced by different stimuli. The detection of sequences of cognitive states can help in the diagnosis of difficulties in performing a complex task. We have implemented a system consisting of two parts: a feature selection component, and a sonification engine. For the feature selection component we investigate different feature selection methods, while for the sonification engine we explore different data to sound mapping strategies. We apply our system to fMRI data produced by auditory stimuli consisting of rhythmic and non-rhythmic audio signals.

The rest of the paper is organized as follows: Section 2 sets out the background for this research. In Section 3, we describe our approach to fMRI data sonification, and finally Section 4 presents some conclusions and indicates some areas of future research.

2. BACKGROUND

2.1. Functional Magnetic Resonance Imaging

Functional Magnetic Resonance Imaging (fMRI) is a brain imaging technique that allows the observation of brain activity in human subjects based on the increase in blood flow to the local vasculature that accompanies neural activity in the brain. More precisely, fMRI measures the ratio of oxygenated hemoglobin to deoxygenated hemoglobin in the blood with respect to a control baseline, at many individual locations within the brain. The blood oxygen level is believed to be influenced by local neural activity, and thus this blood oxygen level dependent (BOLD) response is normally taken as an indicator of neural activity. An fMRI scanner measures the value of the fMRI signal (BOLD response) at all the points in a three dimensional image.

An fMRI scanner produces time-series data that represents brain activity in a collection of 2D slices of the brain. The collection of the 2D slices form a 3D image of the brain containing in the order of 60000 voxels, i.e. cubes of tissue about 2 millimeters on each side. Images are usually taken every 1-5 seconds. Despite the limitations in temporal resolution, fMRI is arguably the best technique for observing human brain activity that is currently available. While the spatial resolution of fMRI is dramatically better than that provided by earlier brain imaging methods, each voxel nevertheless contains on the order of hundreds of thousands of neurons. Figure 1 shows fMRI data collected while a person listened to auditory stimuli.

fMRI has been widely applied to the task of identifying the regions in the brain which are activated when a human performs a particular cognitive function. Most of the reported research summarizes average fMRI responses when a human is presented with a particular stimulus repeatedly. Regions in the brain activated by a particular task are identified by comparing fMRI activity during the period where the stimulus is presented with the activity de-

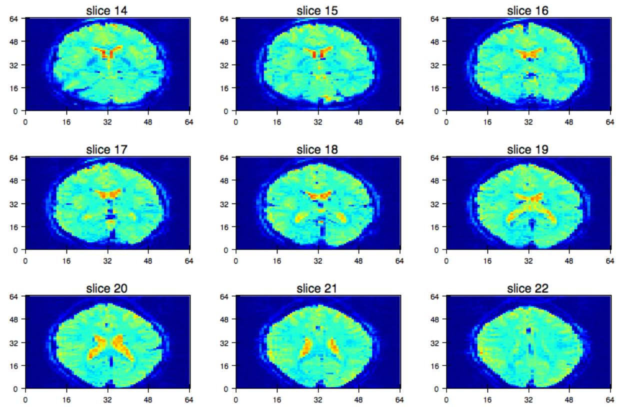

Figure 1: fMRI data collected while a person listened to auditory stimuli. The figure represents nine 2D slices from a 3D image of the brain. Every slice is 64x64 voxels and their intensities are represented with a "jet" colormap . This colormap begins with blue (lowest intensity) , and passes through cyan, yellow, orange, and red (highest intensity).

tected under a control condition. In this paper, the aim is to identify different cognitive states from fMRI data using sonification.

2.2. Sonification

Sonification refers to the use of (non-speech) audio in order to convey information about data. Due to the characteristics of auditory perception, such as excellent temporal and pressure resolution, sonification provides an interesting alternative or complement to visualization techniques. Sonification has been a well established technique in applications that require a constant awareness of some information (e.g. vital body functions during an operation). Success stories of sonification include the Geiger counter, sonar, the auditory thermometer, and numerous medical auditory displays. Recently, several tools have been developed to explore data streams through sonification. This is the case of Sonifyer, a Mac users interface for listening to data, mainly based on audification and FM synthesis [5]. Two other sonification tools are AeSon Toolkit which is motivated by user-centred customisation of the aesthetic representation and scope of the data[6], and SUMO [7] for the sonification of chemical data.

Nowadays, with abundance of high-dimensional data, auditory data exploration has become an important tool to comprehend high-dimensional data and to uncover important structures and patterns [8, 9] in complex data. It is particularly appropriate to improve insight into biomedical data, which are naturally multidimensional. Sonification based on Electroencephalography (EEG) has been widely used for the study of the brain [10, 11, 12, 13].

One of the first attempts to auditory EEG exploration was reported in 1934 by E. Adrian and B. Matthews [10]. They measured the brain activity from a human subject by electrodes applied to the head, and the channels were viewed optically on bromide paper using the Matthews oscillograph, while being directly transduced into sound. They could demonstrate the synchronization between brain activity and external stimuli.

More recently, T. Hermann et al. in 2002 [11] presented different strategies of sonification for human EEG: *spectral mapping*, by analysing the spectral properties of the signals; *Distance Matrix Sonification* , using the Euclidean distance among all signals; and *Differential Sonification*, where they compare the data from different conditions and different channels.

In [14] T. Hermann and G. Baier analysed the rhythmical structure of EEG using auditory exploration. They used a set of differential equations to process the data and extract the parameters to feed the Model-Based Sonification [15]. In 2006 T. Hermann and G. Baier [16] used an articulatory speech model driven by variable features. Both personalized and generic features were used, such as transient activity, spatial distribution or correlation matrix features. T. Hermann and G. Baier also explored multi-channel sonification [13]. The system was intended to allow the listener to perceive spatial characteristics of the data in a multi-speaker environment. They explored the idea of *Event-Based Sonification* (EBS), where features are defined as events that trigger sound synthesis. In this case, local maxima was thought to be suitable both for real-time sonifications and meaningful to the clinician.

There has also been attempts to translate human EEG into music. D. Wu et al. worked to represent mental states by using music [17]. The EEG features were extracted by wavelet analysis and they would control musical parameters such as pitch, tempo, rhythm, and tonality. To give more musically meaning, some rules were taken into account like harmony or structure. One of the main challenges of this work was to find the precise trade-off between directly sonification of the features and music composition.

However, to the best of our knowledge, no similar research projects based on fMRI data have been reported in the scientific literature.

3. THE FMRI SONIFICATION SYSTEM

3.1. Feature Selection

Given the high dimensionality of the data considered, before any attempt to sonification it is necessary to apply feature selection methods. In this paper, we explore the following feature selection strategies:

- **Voxel discriminability**. For each voxel and considered cognitive state, a t-test is performed comparing the fMRI activity of the voxel in examples belonging to the two stimuli of interest. In the case of more than two cognitive states, instead of the t-test, an f-test is performed comparing the fMRI of the voxel in examples belonging to the different stimuli of interest. n voxels are then selected by choosing the ones with larger t-values.
- **Voxel activity**. For each voxel and considered cognitive state, a t-test is performed comparing the fMRI activity of the voxel in examples belonging to a particular stimulus to its activity in examples belonging to fixation periods. For each cognitive state, n voxels are then selected by choosing the ones with larger t-values. Note that these voxels may discriminate only one target class from fixation.

The feature selection strategies are motivated by the fact that fMRI binary cognitive state identification problems naturally give rise to three types of data (similarly for non-binary identification problems): data corresponding to the two target classes, C1 and C2, and data corresponding to the fixation condition. Data corresponding to C1 and C2 is composed of signal plus noise, while data corresponding to the fixation condition contains only noise, i.e. it contains no relevant signal. Thus, two natural feature selection methods are voxel discriminability, i.e. how well the feature discriminates C1 and C2, and voxel activity, i.e. how well the feature distinguishes C1 or C2 from the fixation class. While the former selection method is a straightforward method for selecting voxels which discriminate the two classes, the later focuses on choosing voxels with large signal-to-noise ratios, although it ignores whether the feature actually discriminates the two classes. Within the fMRI community it is common to use voxel activity to select a subset of relevant voxels.

In conjunction with the voxel discriminability and voxel activity strategies, we have explored *Spherical Multivariate Searchlight*. Spherical Multivariate Searchlight is used to obtain a continuous map in which informative regions are marked, by moving a spherical multivariate searchlight through the measured volume of brain activity. The searchlight is centered on each voxel in turn. To combine the signals from all voxels falling into the searchlight, we compute the average of their t-values.

3.2. Voxel sonification

The core of sonification is the processes and algorithms that define the mapping of data to sound for any particular application. The term mapping refers, to mathematical transformations applied to real-time data received from controllers or sensors so that they may be used as effective control for sound synthesis parameters.

For our purpose we have used Parameter-Mapping Sonification, in which data values are mapped to the various parameters of a sound. This approach is particularly suited for multivariate representation, as many data dimensions can be listened to at the same time. Nevertheless, connecting the parameters to components of a real-time sound synthesis system is not trivial.

To effectively perform a musically satisfying mapping, we must understand well the nature of the data sources and the nature of the sounds and music we want to produce. This poses significant problems in the case of biologically controlled sonification in that the goal is to have an unambiguous interpretation of the meaning of biological signals whether direct or derived. Moreover, we should ask ourselves, how should human brain activity sound? how consistent would a sonar interpretation be for various listeners?

The artificial sound synthesis has been implemented by additive synthesis controlled by the features extracted from the data as explained in section 3.1. According to this technique, we have implemented and compared two different sonifications strategies.

In the first approach, every selected feature controls the level of a single note, creating minor blues chords, within several octaves. In order to do that every feature is normalized by its energy activation range to avoid preferences on higher energetic features. However, for each time instant, only the five features with highest activation value will be synthesized. Hence, a singular sound will be created at every instant by means of timbre, pitch and loudness, that represents the activation patterns of the selected features. The intension of this approach is to create harmonic and pleasant sounds. However, the number of extracted features must remain low, limited by the number of octaves that the human auditory system is able to perceive. A sonification sample using this approach can be found at www.upf.dtic.edu/∼rramirez/blues.mp3.

The second sonification strategy uses a larger number of features, approximately 200. The idea is to create a sound texture, that would represent the data by summing partials with additive synthesis. In this case, the normalized energy from every feature is mapped to the frequency of a sine tone within the human hearing range. The resulting sound has a noisy nature due to the fact that there are no harmonic restrictions. Nonetheless, it is the representation of the evolution of the selected features across time. A fragment of the sound spectrogram can be seen in Figure 2 and a sonification sample can be found at www.upf.dtic.edu/∼rramirez/additive.mp3.

The chosen software environment for sonification has been Pure Data [18], since it makes available rapid prototyping of real-time sound generators. It also supports the Open Sound Control protocol for communication between the core and the sound generator. A piece of the sound engine for the first approach can be seen at Figure 3.

Finally, the combination of both visualization and sonification of the data may lead to a better understanding of it. For that purpose, we have implemented an interface, Figure 4, that allows the user to visually explore the data, while hearing the sonification of the selected features, as explained in section 3.1.

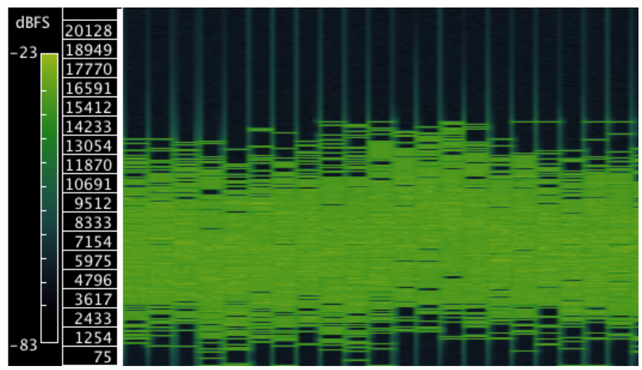

Figure 2: Sound spectrogram fragment from the second sonification approach. The x and y axes represent time and frequency respectively, and the color represents the intensity of the frequency components. The horizontal lines show the partials mapped from the selected features and the vertical lines is the consequence of the abrupt transitions between time slots.

3.3. Experiments and data

The fMRI data used in this study (for details see [19]) was produced by volume accented and duration accented rhythmic stimuli were used. The stimuli were between 14 and 18 s long. There were four rhythm types: volume beat, volume non-beat, duration beat and duration non-beat. Thus, the first rhythm type (Volume accented with Beat) consisted of 81 tones, in which every 4th tone was louder by 6.6 dB, in order to give rise to the perception of a regular beat (occurring 21 times per trial). For each trial, the tone length was chosen from a range of 180 to 228 ms (in 8 ms steps) so that a new beat would be induced in each trial, not simply carried over from a previous trial. Accordingly, the beat occurred at a rate of 720 to 912 ms. The second rhythm type (Volume accented with no beat) also had 81 tones. However, the tone volumes were not isochronous, so no regular beat could be fit to the rhythm.

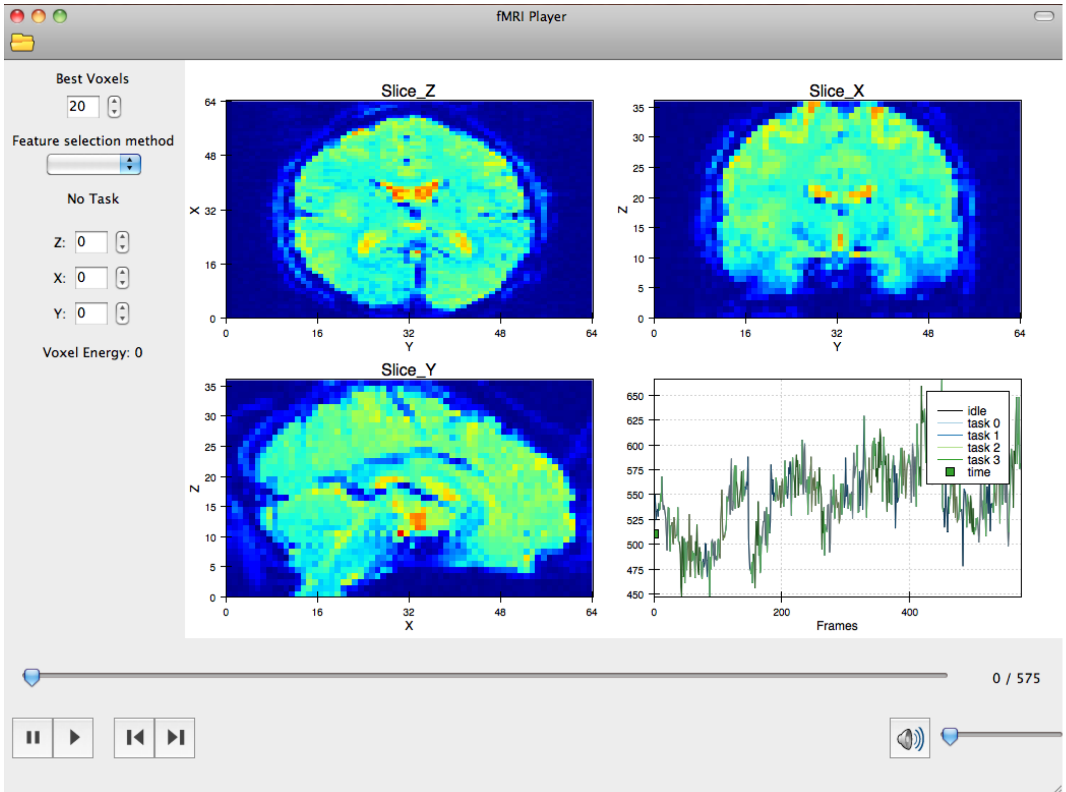

Figure 4: fMRI Data Interface. The plots at the top and at the bottom left represent three different slices of a 3D image of the brain, according to the X,Y, Z axes. The fourth plot represents the energy of the voxel selected by the user, across time. The bottom panel is used a video player and the left panel shows different information about the system and the selected voxel.

4. CONCLUSIONS AND FUTURE WORK

We have proposed two different fMRI data sonification approaches to cognitive state identification. The first approach brings a harmonic sonification to explore the data by using blues chords as reference. The second approach creates a complex texture sound by using a large amount of features coming from the data.

The system's objective is the auditory detection of cognitive states produced by different auditory stimuli, and combines sonification and visualization in order to incorporate the benefits of both techniques. We have explored different feature selection techniques in order to reduce the dimensionality of the date before sonification. In particular we have explored voxel discriminability and voxel activity feature selection. The work reported is still in progress but the results we have so far obtained are encouraging. This preliminary results seem to indicate that the fMRI data considered contain sufficient information to identify different cognitive states by sonifying a small number of features (i.e. 20 voxels) extracted from the studied fMRI data, and with no prior anatomical knowledge. The problem provides a very interesting instance of sonification with extremely high dimensional, sparse and noisy data. As future work, we plan to explore additional feature extraction methods and to conduct a series of experiments for quantitatively evaluating the system.

Acknowledgements. This work was supported by the Spanish Ministry of Science and Innovation under grant TIN2009-14247-C02-01 DRIMS Project and the "Obra Social la Caixa" grant for master degree studies. We would like to thank Jessica Grahn for providing the fMRI data used in this paper.

5. REFERENCES

[1] T. Hermann, P. Meinicke, H. Bekel, H. Ritter, H. M. Mueller, and S. Weiss, "Sonifications for eeg data analysis," R. Nakatsu and H. Kawahara, Eds., Advanced Telecommunications Research Institute (ATR), Kyoto, Japan. Kyoto, Japan: Advanced Telecommunications Research Institute (ATR), Kyoto, Japan, 2002. [Online]. Available: Proceedings/2002/HermannMeinicke2002.pdf

[2] T. Kohonen, "The self-organizing map," *Proceedings of the IEEE*, vol. 78, no. 9, pp. 1464 –1480, Sep. 1990.

[3] S. T. Roweis and L. K. Saul, "Nonlinear dimensionality reduction by locally linear embedding." *Science (New York, N.Y.)*, vol. 290, no. 5500, pp. 2323–6, Dec. 2000. [Online]. Available: http://www.ncbi.nlm.nih.gov/pubmed/11125150

[4] J. Friedman and J. Tukey, "A projection pursuit algorithm for exploratory data analysis," *Computers, IEEE Transactions on*, vol. C-23, no. 9, pp. 881 – 890, 1974.

[5] F. Dombois, "Sonifyer a concept, a software, a platform," Paris, France, 2008, inproceedings. [Online]. Available: Proceedings/2008/Dombois2008.pdf

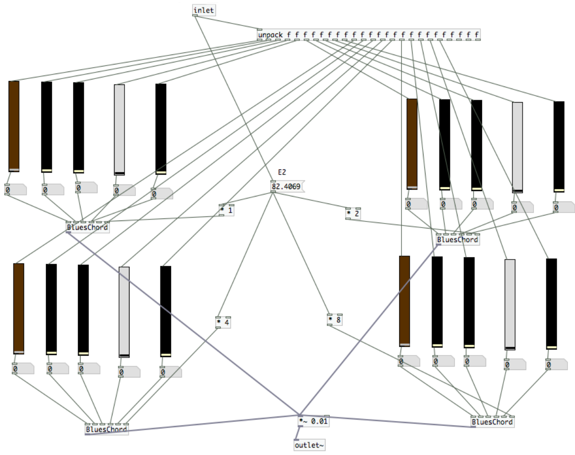

Figure 3: Pure Data sound generator. The sliders are connected to the fMRI selected features and represent the sound level of single note within blues minor chord. The groups of five sliders represent different octaves of the same chord.

[6] K. Beilharz and S. Ferguson, "An interface and framework design for interactive aesthetic sonification," Copenhagen, Denmark, 18-21 May 2009. [Online]. Available: Proceedings/2009/BeilharzFerguson2009.pdf

[7] F. Grond and F. Dall'Antonia, "Sumo. a sonification utility for molecules," Paris, France, 2008, inproceedings. [Online]. Available: Proceedings/2008/GrondDallAntonia2008.pdf

[8] T. Hermann, M. Hansen, and H. Ritter, "Combining Visual and Auditory Data Exploration for finding structure in high-dimensional data," *sonification.de*, pp. 1–10. [Online]. Available: http://www.sonification.de/publications/media/HermannHansenRitter2001-CVA.pdf

[9] S. Barrass and G. Kramer, "Using sonification," *Multimedia Systems*, vol. 7, no. 1, pp. 23–31, Jan. 1999. [Online]. Available: http://www.springerlink.com/openurl.asp?genre=article\&id=doi:10.1007/s005300050108

[10] E. D. Adrian and B. H. C. Matthews, "The Berger Rhythm: potential changes from the occipital lobes in man," *Brain*, vol. 57, pp. 355–384, 1934.

[11] T. Hermann, P. Meinicke, H. Bekel, H. Ritter, H. M. Mueller, and S. Weiss, "Sonifications for eeg data analysis," R. Nakatsu and H. Kawahara, Eds., Advanced Telecommunications Research Institute (ATR), Kyoto, Japan. Kyoto, Japan: Advanced Telecommunications Research Institute (ATR), Kyoto, Japan, 2002. [Online]. Available: Proceedings/2002/HermannMeinicke2002.pdf

[12] B. D. Sonification, T. Hermann, and G. Baier, "The sonification of human electroencephalogram Outline Audification of EEG Audification in practise ! Load Data into Buffers," pp. 1–9, 2006.

[13] G. Baier, T. Hermann, and U. Stephani, "Multi-channel sonification of human eeg," G. P. Scavone, Ed., Schulich School of Music, McGill University. Montreal, Canada: Schulich School of Music, McGill University, 2007, pp. 491–496. [Online]. Available: Proceedings/2007/BaierHermann2007.pdf

[14] G. Baier and T. Hermann, "The sonification of rhythms in human electroencephalogram," S. Barrass and P. Vickers, Eds., International Community for Auditory Display (ICAD). Sydney, Australia: International Community for Auditory Display (ICAD), 2004. [Online]. Available: Proceedings/2004/BiaerHermann2004.pdf

[15] T. Hermann and H. Ritter, "Listen to your data: Model-based sonification for data analysis," *Advances in intelligent computing and multimedia systems, Baden-Baden, Germany*, pp. 189–194, 1999. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.42.7970\&rep=rep1\&type=pdf

[16] T. Hermann, G. Baier, U. Stephani, and H. Ritter, "Vocal sonification of pathologic eeg features," C. F. A. D. N. E. Tony Stockman, Louise Valgerur Nickerson and D. Brock, Eds., Department of Computer Science, Queen Mary, University of London, UK. London, UK: Department of Computer Science, Queen Mary, University of London, UK, 2006, pp. 158–163. [Online]. Available: Proceedings/2006/HermannBaier2006.pdf

[17] D. Wu, C. Li, Y. Yin, C. Zhou, and D. Yao, "Music composition from the brain signal: representing the mental state by music." *Computational intelligence and neuroscience*, 2010. [Online]. Available: http://dx.doi.org/10.1155/2010/267671

[18] M. Puckette, "Pure Data: another integrated computer music environment," in *Proc. the Second Intercollege Computer Music Concerts*, 1996, pp. 37–41.

[19] J. Grahn and J. Rowe, "Feeling the beat: premotor and striatal interactions in musicians and non-musicians during beat perception," 2009, pp. 7540–7548.